(NASA-CR-175504) A CONCATENATED CODING SCHEME FOR ERROR CONTROL (Hawaii Univ., Manoa.) 12 p HC A02/MF A01 CSCL 09B

N85-22001

Unclas

G3/61 14486

A CONCATENATED CODING SCHEME FOR ERROR CONTROL

Technical Report

to

NASA
Goddard Space Flight Center
Greenbelt, Maryland

Grant Number NAG 5-407

Shu Lin
Principal Investigator
Department of Electrical Engineering
University of Hawaii at Manoa
Honolulu, Hawaii 96822

March 20, 1985

A CONCATENATED CODING SCHEME FOR ERROR CONTROL

ABSTRACT

In this paper, a concatenated coding scheme for error control in data communications is analyzed. In this scheme, the inner code is used for both error correction and detection, however the outer code is used only for error detection. A retransmission is requested if the outer code detects the presence of errors after the inner code decoding. In this paper, the probability of undetected error of the above error control scheme is derived and upper bounded. Two specific example schemes are analyzed. In the first example scheme, the inner code is a distance-4 shortened Hamming code with generator polynomial $(X+1)(X^6+X+1) = X^7+X^6+X^2+1$ and the outer code is a distance-4 shortened Hamming code with generator polynomial $(X+1)(X^{15}+X^{14}+X^{13}+X^{12}+X^4+X^3+X^2+X+1) = X^{16}+X^{12}+X^5+1$ which is the X.25 standard for packet-switched data network. This example scheme is proposed for error control on NASA telecommand links. In the second example scheme, the inner code is the same as that in the first example scheme but the outer code is a shortened Reed-Solomon code with symbols from $GF(2^8)$ and generator polynomial $(X+1)(X+\alpha)$ where α is a primitive element in $GF(2^8)$. We show that both example schemes provide very high reliability.

1. Introduction

Consider a concatenated coding scheme for error control for a binary symmetric channel with bit-error-rate $\varepsilon<1/2$ as shown in Figure 1. Two linear block codes, C_f and C_b, are used. The inner code C_f, called frame code, is an (n,k) code with minimum distance d_f. The frame code is designed to correct t or fewer errors and simultaneously detect $\lambda(\lambda\geq t)$ or fewer errors where $t+\lambda+1\leq d_f$. The outer code C_b is an (n_b,k_b) code with minimum distance d_b and $n_b=mk$, where m is a positive integer. The outer code is designed for error detection only.

The encoding is done in two stages. A message of k_b bits is first encoded into a codeword of n_b bits in the outer code C_b. Then the n_b-bit word is divided into m k-bit segments. Each k-bit segment is encoded into an n-bit word in the frame code C_f. This n-bit word is called a frame. Thus, corresponding to each k_b-bit message at the input of the outer code encoder, the output of the frame code encoder is a sequence of m frames. This sequence of m frames is called a block. A two dimensional block format is depicted in Figure 2.

The decoding consists of error correction in frames and error detection in m decoded k-bit segments. When a frame in a block is received, it is decoded based on the frame code C_f. The $n-k$ parity bits are then removed from the decoded frame, the k-bit decoded segment is stored in a buffer. If there are t or fewer transmission errors in a received frame, the errors will be corrected and the decoded segment is error free. If there are more than λ errors in a received frame, the decoded segment may contain undetected errors. After m frames of a block have been decoded, the buffer contains m k-bit decoded segments. Then error detection is performed on these m decoded segments based on the outer code C_b. If no error is detected, the m decoded segments are assumed to be error free and are accepted by the receiver. If the presence of errors is detected, the m decoded segments are discarded and the receiver requests a retransmission of the rejected block. Retransmission and decoding process continues until a transmitted block is successfully received. Note that a successfully received block may be either error free or contains undetectable errors.

The error control scheme described above is actually a combination of forward-error-correction (FEC) and automatic-repeat-request (ARQ), called a hybrid ARQ scheme [1]. The retransmission strategy determines the system throughput, it may be one of the three basic modes namely, stop-and-wait, go-back-N or selective-repeat. The reliability is measured in terms of the probability of undetected error after decoding.

In this paper, the probability of undetected error of the above error control scheme is derived and upper bounded. Two specific example schemes are analyzed. In the first example scheme, the inner code is a distance-4 shortened Hamming code with generator polynomial $(X+1)(X^6+X+1) = X^7+X^6+X^2+1$ and the outer code is a distance-4 shortened Hamming code with generator polynomial $(X+1)(X^{15}+X^{14}+X^{13}+X^{12}+X^4+X^3+X^2+X+1) = X^{16}+X^{12}+X^5+1$ which is the X.25 standard for packet-switched data network. This example scheme is proposed for error control on NASA telecommand links. In the second example scheme, the inner code is the same as that in the first example scheme but the outer code is a shortened Reed-Solomon code with symbols from $GF(2^8)$ and generator polynomial $(X+1)(X+\alpha)$ where α is a primitive element in $GF(2^8)$. We show that both example schemes provide very high reliability.

2. Probability of Undetected Error

The probability $P_f(\bar{e}_0,\varepsilon)$ that a decoded frame contains a nonzero error vector $\bar{e}_0$ after decoding is given by [2,3,4],

$$P_f(\bar{e}_0,\varepsilon) = \sum_{i=0}^{t} \sum_{j=0}^{\min(t-i,n-w)} \binom{w}{i}\binom{n-w}{j} \varepsilon^{w-i+j}(1-\varepsilon)^{n-w+i-j} \tag{1}$$

where w is the weight of $\bar{e}_0$. The right-hand side of (1) only depends on w, t and ε, we denote the right-hand side of (1) as $Q_t(w,\varepsilon)$.

Recall that a codeword in the outer code C_b consists of m k-bit segments. At the receiver, error detection is performed on every m decoded segments based on C_b. Let $P_b(\bar{e},\varepsilon)$ denote the probability that the decoded word contains an undetectable error pattern $\bar{e}$ (a nonzero codeword in C_b). For a codeword $\bar{v}$ in C_b, let $\bar{v}^{(j)}$ denote the j-th segment of $\bar{v}$, and let $w_j(\bar{v})$ be the weight of the codeword in frame code C_f into which $\bar{v}^{(j)}$ is encoded. Then it follows from (1) that for an undetectable error pattern $\bar{e}$ in a block

$$P_b(\bar{e},\varepsilon) = \prod_{j=1}^{m} Q_t(w_j(\bar{e}),\varepsilon) \ . \tag{2}$$

Let $P_{ud}^{(b)}(\varepsilon)$ be the probability of undetected error for the outer code C_b. Then

$$P_{ud}^{(b)}(\varepsilon) = \sum_{\bar{e}\varepsilon C_b-\{\bar{0}\}} P_b(\bar{e},\varepsilon) \ . \tag{3}$$

For $1\le j_1<j_2<\dots<j_h\le m$, consider the set of codewords in C_b where nonzero bits are confined in the j_1-th segment, the j_2-th segment, ..., and the j_h-th segment. This set of codewords forms a subcode of C_b, call a $(j_1,j_2,\dots,j_h)$-subcode of C_b and denoted by $C_b(j_1,j_2,\dots,j_h)$. If C_b is a cyclic or shortened cyclic code, then

1. for h=1, all (j_1)-subcodes of C_b are equivalent;
2. for $h\ge 2$, all $(j_1,j_2,\dots,j_h)$-subcodes of C_b with the same $j_2-j_1,j_3-j_2,\dots,j_h-j_{h-1}$ are equivalent codes and are called h-segment $(j_2-j_1,j_3-j_2,\dots,j_h-j_{h-1})$ subcodes of C_b.

Consider a $(j_1,j_2,\dots,j_h)$-subcode of C_b. Let $i_1,i_2,\dots,i_h,r_1,r_2,\dots,r_h$ be a set of integers for which $0\le i_q\le k$ and $0\le r_q\le n-k$ with $1\le q\le h$. Let

$A^{j_1,j_2,\ldots,j_h}_{(i_1,r_1)(i_2,r_2)\ldots(i_h,r_h)}$ denote the number of codewords $\bar{v}$ in $C_b(j_1,j_2,\ldots,j_h)$ such that, for $1\le q\le h$, the j_q-th segment $\bar{v}^{(j_q)}$ of $\bar{v}$ has weight i_q and $w_{j_q}(\bar{v}) = i_q+r_q$. Then it follows from (2), (3) and the definition of $A^{j_1,j_2,\ldots,j_h}_{(i_1,r_1)(i_2,r_2)\ldots(i_h,r_h)}$ that

$$P^{(b)}_{ud}(\varepsilon) = \sum_{h=1}^{m} Q_1(0,\varepsilon)^{m-h}\{ \sum_{1\le j_1<j_2<\ldots<j_h\le m} \sum_{IR_h} A^{j_1,j_2,\ldots,j_h}_{(i_1,r_1)(i_2,r_2)\ldots(i_h,r_h)} \prod_{q=1}^{h} Q_t(i_q+r_q,\varepsilon)\} , \quad (4)$$

where

$$IR_h = \{((i_1,r_1),(i_2,r_2),\ldots,(i_h,r_h)):\ 1\le i_q\le k,\ 0\le r_q\le n-k,\ d_f\le i_q+r_q\ (1\le q\le h)$$
$$\text{and } d_b\le \sum_{q=1}^{h} i_q\le n_b \quad .$$

If C_b is a cyclic or shortened cyclic code, then Eq. (4) can be simplified as follows:

$$P^{(b)}_{ud}(\varepsilon) = \sum_{h=1}^{m} Q_t(0,\varepsilon)^{m-h}\{ \sum_{1\le j_1<j_2<\ldots<j_g\le m} \sum_{IR_h} A^{j_1,j_2,\ldots,j_h}_{(i_1,r_1)(i_2,r_2)\ldots(i_h,r_h)} \prod_{q-1}^{h} Q_t(i_q+r_q,\varepsilon) \quad , \quad (5)$$

From (4) we see that, if we know the detail weight structure of $C_b(j_1,j_2,\ldots,j_h)$, the error probability $P^{(b)}_{ud}(\varepsilon)$ can be computed. However, for a given C_b, it is not easy to find $A^{j_1,j_2,\ldots,j_h}_{(i_1,r_1)(i_2,r_2)\ldots(i_h,r_h)}$. To overcome this difficulty, we have derived upper bounds on the terms on the right-hand side of (5) [5]. We assume that $\varepsilon \le (t+2)/(3t+4)$. Suppose that t=1 and the inner code is an even-weight code and the outer code is a cyclic or shortened cyclic even-weight code. Let $\{A^{(b)}\}$ be the weight distribution of the outer code C_b. We have obtained the following bound on $P_{ud}(\varepsilon)$:

$$
\begin{aligned}
P_{ud}^{(b)}(\varepsilon) \le\; & m \sum_{i=d_b}^{10} \sum_{r=0}^{n-k} A^{1}_{(i,r)} Q_1(i+r,\varepsilon) \\
& + \sum_{j=2}^{m} (m-j+1) \sum_{\substack{i_1, i_2 \le 10 \\ 1 \le i_1, i_2}} A^{1,j}_{i_1,i_2} \prod_{p=1}^{2} Q_1(\beta(i_p),\varepsilon) \\
& + \{ \sum_{i=d_b}^{10} (A_i^{(b)} - mA_i^{1}) - \sum_{j=2}^{m} (m-j+1) \sum_{\substack{i_1, i_2 \le 10 \\ i \le i_1, i_2}} A^{i,j}_{i_1,i_2} \} Q_1(4,\varepsilon)^3 \\
& + \min\{ \binom{m}{3} \binom{k}{4}^2 \binom{k}{3},\; A_{12}^{(b)} \} Q_1(4,\varepsilon)^3 + A_{12}^{(b)} Q_1(6,\varepsilon) \\
& + \sum_{i=4}^{6} A_{4i}^{(b)} Q_1(4,\varepsilon)^i + \sum_{i=3}^{5} A_{4i+2}^{(b)} Q_1(4,\varepsilon)^{i-1} Q_1(6,\varepsilon) \\
& + (26/n_b)^{-26} (1-26/n_b)^{n_b-26} Q_1(4,\varepsilon)^5 Q_1(6,\varepsilon)
\end{aligned}
\tag{6}
$$

where

$$
A^{j_1,j_2,\ldots,j_h}_{i_1,i_2,\ldots,i_h} = \sum_{r_1=0}^{n-k} \sum_{r_2=0}^{n-k} \cdots \sum_{r_h=0}^{n-k} A^{j_1,j_2,\ldots,j_h}_{(i_1,r_1)(i_2,r_2)\ldots(i_h,r_h)} , \tag{7}
$$

and

$$
\beta(i) = \begin{cases} d_f, & \text{for } i \le d_f \\ i, & \text{for even } i \text{ and } i > d_f \\ i+1, & \text{otherwise.} \end{cases}
$$

On the other hand, it follows from (5) that

$$
P_{ud}^{(b)}(\varepsilon) \ge m\, Q_1(0,\varepsilon)^{m-1} \sum_{i=d_b}^{10} \sum_{r=0}^{n-k} A^{1}_{(i,r)} Q_1(i+r,\varepsilon) . \tag{9}
$$

3. Examples

We consider two examples of the concatenated coding scheme.

Example 1: The frame code C_f is a distance-4 Hamming code with generator polynomial,

$$
\bar{g}_f(X) = (X+1)(X^6+X+1) = X^7+X^6+X^2+1 ,
$$

where X^6+X+1 is a primitive polynomial of degree 6. The maximum length of this code is 63. This code is used for single error correction. The code is capable of detecting all the error patterns of double and odd number errors. The outer code is also a distance-4 shortened Hamming code with generator polynomial,

$$\bar{g}_0(X) = (X+1)X^{15}+X^{14}+X^{13}+X^{12}+X^4+X^3+X^2+X+1) = X^{16}+X^{12}+X^5+1 ,$$

where $X^{15}+X^{14}+X^{13}+X^{12}+X^4+X^3+X^2+X+1$ is a primitive polynomial of degree 15. We assume that the number of frames in a block is greater than 3 and less than 65. The 16 parity bits of this code is used for error detection only. This scheme is proposed for NASA telecommand system. For this example upper bounds on the probability of undetected error have been computed in [5].

Example 2: This example is a variation of example 1. The frame code C_f is the same as example 1. The outer code is a shortened Reed-Solomon code with generator polynomial $(X+1)(X+\alpha)$ over $GF(2^8)$, where α is a root of $X^8+X^4+X^3+X^2+1$.

For various ε, k, and m, the bound on $P_{ud}^{(b)}(\varepsilon)$ given by (6) is evaluated and plotted in Figures 3 to 5.

ACKNOWLEDGMENT

The authors would like to thank Mr. Toyoo Takata for his helpful support.

REFERENCES

1. S. Lin and D.J. Costello, Jr., Error Control Coding: Fundamentals and Applications, Prentice-Hall, New Jersey, 1983.

2. E.R. Berlekamp, Algebraic Coding Theory, McGraw-Hill, New York, 1968.

3. F.J. MacWilliams, "A Theorem on the Distribution of Weights in Systematic Code," Bell System Technical Journal, Vol. 42, pp. 79-94, 1963.

4. Z. McHuntoon and A.M. Michelson: "On the Computation of the Probability of Post-Decoding Error Events for Block Codes," IEEE Trans. on Information Theory, Vol. IT-23, No. 3, May 1977, pp. 399-403.

5. A. Kitai, "A Method for Computing Probability of Undetectable Error of Error Correcting Code," Thesis for M.E. Degree, Dept. of Information and Computer Sciences, Osaka University, 1984.

6. D.J. Costello, Jr., private communication.

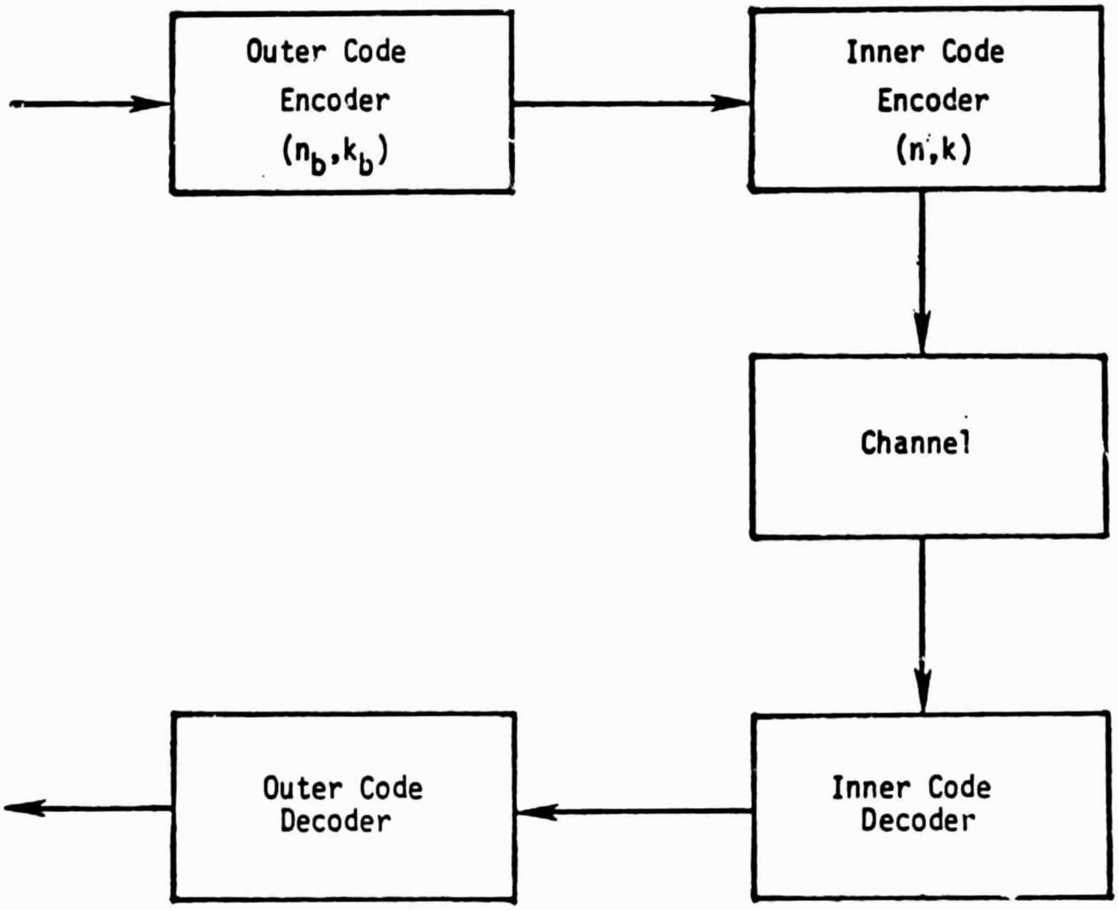

Figure 1 A concatenated coding scheme

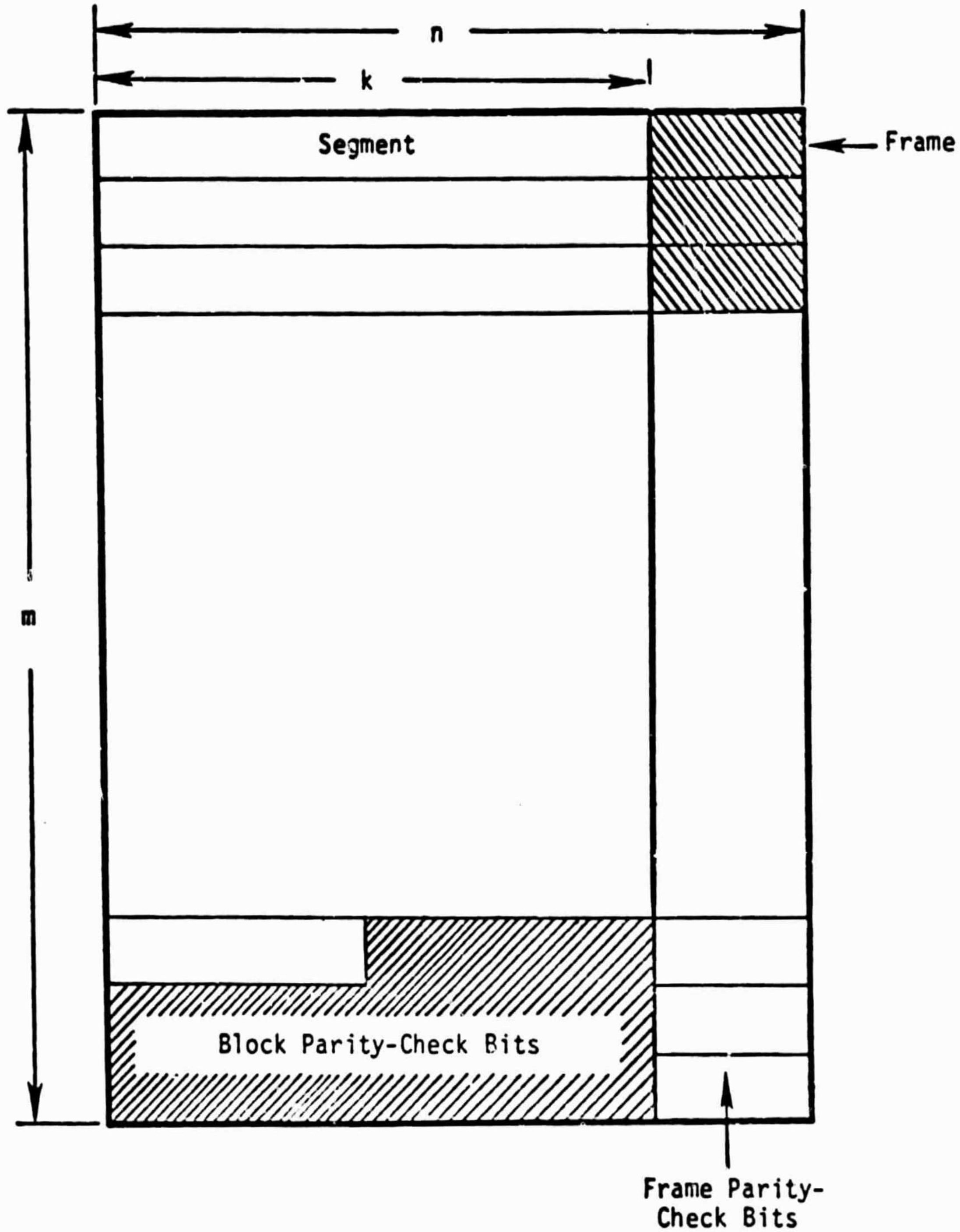

Figure 2 Block format

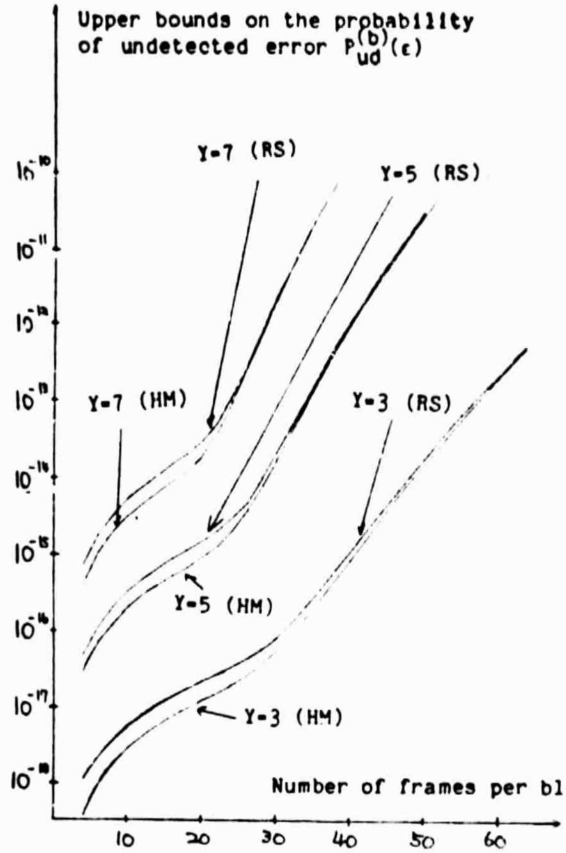

Figure 3 Upper bounds on the probability of undetected error for bit error rate $\varepsilon = 10^{-4}$.

Y: the number of information bytes in a frame

HM: example 1 RS: example 2

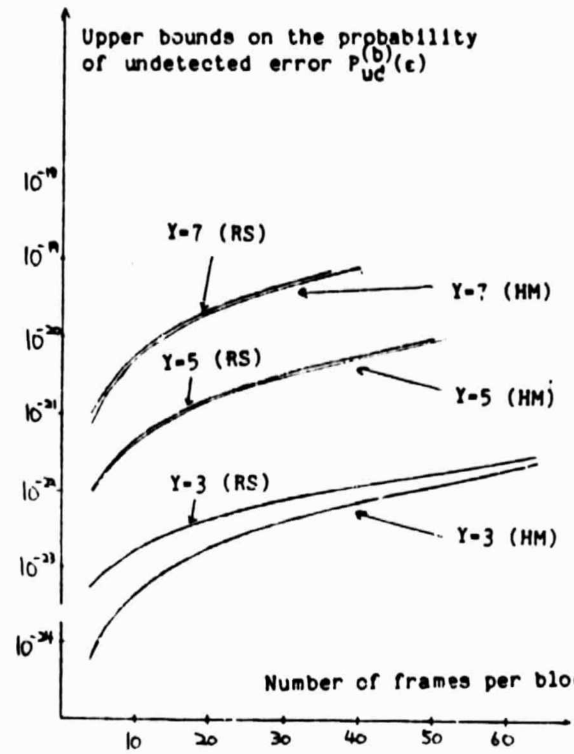

Figure 4 Upper bounds on the probability of undetected error for bit error rate $\varepsilon = 10^{-5}$.

Y: the number of information bytes in a frame

HM: example 1 RS: example 2

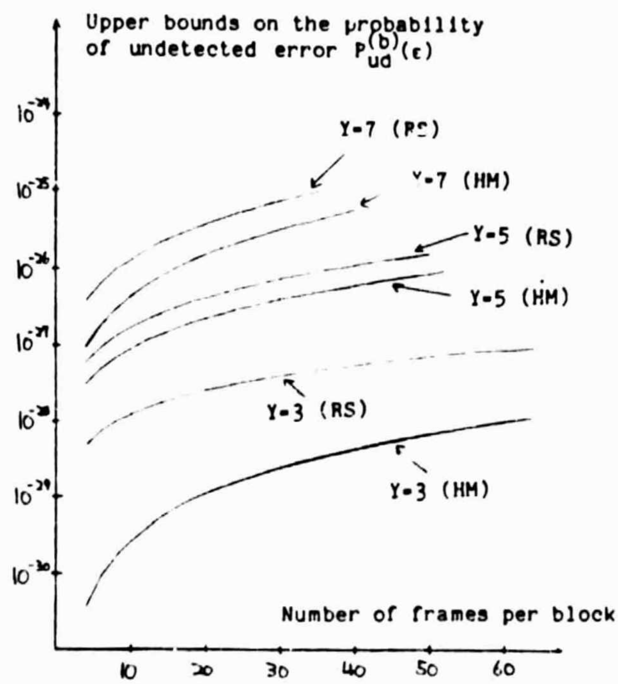

Figure 5 Upper bounds on the probability of undetected error rate $\varepsilon = 10^{-6}$.

Y: the number of information bytes in a frame

HM: example 1 RS: example 2